垫状卷柏 *SpLEA*1 基因克隆及干旱胁迫下的表达分析

周璇，高鹏华，鄢波*

（西南林业大学 园林园艺学院，昆明 650224）

摘要：晚期胚胎发育丰富蛋白（late embryogenesis abundant，LEA）与植物抗逆性密切相关，在干旱胁迫下保护植物细胞，减少植物损伤，并广泛存在于生物体内。垫状卷柏（*Selaginella pulvinata*）是一种在干旱胁迫下生存能力极强的蕨类植物，具有很强的恢复能力。为探究垫状卷柏 *SpLEA*1 基因在耐旱植物中的分子机制与表达特征，该研究以高耐旱性植物垫状卷柏为实验材料，基于转录组测序结果，采用 RT-PCR 技术获得 *SpLEA*1 基因 cDNA 序列，采用 HiTail-PCR 技术获得启动子序列，利用生物信息学对序列进行了分析，采用 qRT-PCR 技术，分析了 *SpLEA*1 基因在干旱胁迫下的表达模式。结果表明：（1）垫状卷柏 *SpLEA*1 全长为 476 bp，开放阅读框（ORF）为 279 bp，共编码 92 个氨基酸，通过在线工具预测到蛋白分子量为 9 491.46 Da, 等电点为 5.45，蛋白结构预测分析表明该蛋白为亲水性蛋白，含有 10 个磷酸化位点，二级结构以 α-螺旋和无规则卷曲为主。（2）预测到 SpLEA1 蛋白的保守结构域为 Lea-5，来源于 LEA1 家族，基于系统发生树和遗传距离矩阵，发现垫状卷柏 SpLEA1 与来自鹰嘴豆（*Cicer arietinum* ）和红三叶（*Trifolium pratense*）的 Lea-5 蛋白同源性较高。（3）利用在线工具对启动子序列进行顺式作用元件的预测分析发现 *SpLEA*1 基因启动子含有 5 类激素响应元件和与干旱胁迫响应有关的功能元件。（4）在自然干旱处理下 *SpLEA*1 基因表达上调并在 12 h 时达到峰值，在 24 h 干旱后进行复水处理，表达量显著下调。综上所述，*SpLEA*1 基因在垫状卷柏中很可能参与了干旱胁迫响应机制的相关调控。此结果为进一步研究垫状卷柏 *SpLEA*1 基因在干旱胁迫下的功能及其表达调控机制奠定了基础。

关键词：垫状卷柏，*SpLEA*1，基因克隆，启动子克隆，表达分析

中图分类号：Q943　　　　　　　　**文献标识码：A**

Cloning and expression analysis of the *SpLEA*1 gene of *Selaginella pulvinata* under drought stress

ZHOU Xuan, GAO Penghua, YAN Bo*

(*College of Landscape Architecture and Horticulture Sciences*, *Southwest Forestry University*, Kunming 650224)

Abstract: Late embryogenesis abundant (LEA) is closely related to plant resilience, protects plant cells, reduces plant damage under drought stress, and is widely present in organisms. LEA proteins are less studied in ferns, and *Selaginella pulvinata* is a fern with the ability to survive drought stress, and it is a fern with a strong recovery ability under drought stress. To investigate the molecular mechanisms and expression characteristics of the *SpLEA*1 gene in drought-tolerant plants, this study used the highly drought-tolerant plant *S. pulvinata* as experimental material and obtained the cDNA sequence of the *SpLEA*1 gene by RT-PCR based on the transcriptome sequencing results. The promoter sequence was obtained by the HiTail-PCR technique, and the

基金项目：国家自然科学基金（31160177） [Supported by National Natural Science Foundation of China (31160177)]。

第一作者：周璇（1996-），硕士，主要从事植物分子生物学研究，(E-mail)974865841@qq.com。

***通信作者：**鄢 波，博士，教授，博士研究生导师，主要从事分子生物方向研究，(E-mail)yanbodr@aliyun.com。

sequence was analyzed by bioinformatics. qRT-PCR was used to analyze the expression pattern of the *SpLEA*1 gene under drought stress. The results were as follows: (1) The length of *SpLEA*1 was 476 bp, the open reading frame (ORF) was 279 bp, and it encoded 92 amino acids. The predicted molecular weight of the protein was 9 491.46 Da, and the isoelectric point was 5.45. The predicted protein structure analysis showed that the protein was hydrophilic. The protein contained 10 phosphorylation sites, of which 6 serine, 3 tyrosine, and 1 threonine, respectively, and the predicted secondary and tertiary structures showed that the protein was mainly composed of α-helix and random coil. (2) The conserved structural domain of the SpLEA1 protein was predicted to be Lea-5, derived from the LEA1 family. Based on the phylogenetic tree and genetic distanced matrix, the SpLEA1 was found to have high homology with Lea-5 protein from *Cicer arietinum* and *Trifolium pratense*. (3) Predictive analysis of *cis*-acting elements in promoter sequenced using online tools revealed that the *SpLEA*1 gene promoter contained five classes of hormone response elements and functional elements related to the drought stress response. The *SpLEA*1 gene was hypothesized to have multiple functions in the plant body and was closely related to drought stress response mechanisms. (4) *SpLEA*1 gene expression was up-regulated under natural dehydration treatment and peaked at 12 h. After rehydration treatment at 24 h, expression was significantly down-regulated. In summary, the *SpLEA*1 gene is likely to be involved in the regulation of drought stress response mechanisms in matted curly cypress. This result lays the foundation for further studies on the function of the matted cypress *SpLEA*1 gene under drought stress and its expression regulation mechanism.

Keywords: *Selaginella pulvinata*, *SpLEA*1, genetic cloning, promoter cloning, expression analysis

植物在长期进化过程中，对外界的生物胁迫与非生物胁迫有系统的生理及分子响应机理。研究报道指出，胚胎发育晚期丰富蛋白 (late embryogenesis abundant，LEA)与植物抗逆性相关，广泛参与植物对非生物胁迫的响应过程（李翔，2016）。LEA 蛋白最初是在棉花种子中分离克隆得到的（Dure et al., 1981），之后研究发现 LEA 蛋白广泛存在于植物、无脊椎动物与原核生物中。*LEA* 基因在植物的整个发育阶段均有表达，特别是在植物遭受如干旱，高温等环境胁迫时，*LEA* 基因在植物组织细胞中大量表达，积累丰富的 LEA 蛋白以应对外界环境（Wise, 2003；Silveira et al., 2008）。2010 年，Hunault & Jaspard（2010）建立了 LEA 蛋白数据库(Late Embryogenesis Abundant Proteins database，LEAPdb)，根据 LEA 蛋白氨基酸序列的 8 个保守的 PFAM 结构域对 LEA 蛋白成员进行生物信息学分析。通过 LEA 蛋白数据库确认研究中所提取的 LEA 蛋白的 PFAM 号，可以确定其所属的 LEA 蛋白家族，为后续实验提供基础。 LEA1(group 1 late-embryogenesis-abundant proteins)是以无规则结构形式存在的亲水性蛋白，在植物中广泛分布，典型代表为棉花 D-19、小麦 EM 蛋白、大麦 B19 蛋白等。LEA1 家族成员均具有亲水性，各成员之间具有的亲水性极高的 20 个氨基酸的保守基序（GGETRKEQLGEEGYREMGRK）数量多变（Stacy et al., 1995）。

在植物中克隆目的基因的启动子能进一步系统分析基因的功能。植物启动子是一段含有转录起始位点，调控基因表达的 DNA 序列，启动子的转录频率、起始方向和位点均是基因转录调控表达的关键（Liu et al., 1997；王志等，2011；张曦予等，2019）。启动子主要包含一些特异性的调控基序（梅玉芹，2018），在结构和功能上可以分为组成型启动子、诱导型启动子和组织特异型启动子（杨瑞娟等，2018）。研究表明，诱导型启动子在外界环境因素发生改变时，会使基因瞬时或持续性上调表达（D'urzo et al., 2013）。Zheng 等（2019）从卷心菜中分离得到了一个非典型 *LEA* 基因 *LpLEA* 基因及其启动子序列，分析表明 *LpLEA* 基因的启动子存在与非生物胁迫有关的独特顺式作用调控元件，*LpLEA* 基因在不同非生物胁迫

及脱落酸诱导下在不同部位表达量均提高。

垫状卷柏（*Selaginella pulvinata*）又名九死还魂草，主要分布于我国的干旱地区，常生长于裸露的石灰岩表面或者石缝中，具有很强的耐旱性，属于蕨类植物门垫状卷柏科垫状卷柏属土生或石生的复苏植物（吴征镒，2004）。研究表明，垫状卷柏具有独特的活性氧生成和清除调节途径，增强脱落酸生物合成和潜在的调节脱落酸信号及对脱落酸响应的机制，且对其叶绿体基因组分析发现其叶绿体结构具有独特的重排且叶绿体 NAD（P）H 脱氢酶（NDH）基因完全缺失（Saucedo et al., 2017）。 LEA1 蛋白可在植物幼苗时期受干旱、盐胁迫、ABA 以及低温胁迫诱导表达，且对植物体内乳酸脱氢酶的活性有保护作用，同时可正向调控部分钙依赖性蛋白激酶的表达（邹永东，2011；Xiang et al., 2018），LEA1 蛋白作为参与植物耐受性调控的重要蛋白在高耐旱植物垫状卷柏中发挥的作用值得探究。目前，有关蕨类植物 LEA 蛋白的研究较少，在蕨类植物垫状卷柏中 *LEA*1 基因的研究几乎空白。因此，在垫状卷柏中克隆 *LEA*1 基因并分析其分子机制及表达特征，对深入研究其在垫状卷柏抗旱过程中的调控机制具有重要意义。本研究分离克隆了垫状卷柏 *SpLEA*1 基因，对 *SpLEA*1 基因序列及启动子顺式作用元件进行了生物信息学分析，构建了系统发生树和遗传距离矩阵，对同源蛋白序列进行了比对，利用实时荧光定量技术检测了垫状卷柏幼嫩叶片不同干旱状态下的表达情况，为进一步探索 *SpLEA*1 基因在干旱胁迫下的功能及分子作用机制奠定基础，同时，为园林园艺观赏植物的抗旱性改良提供基因资源。

1　材料与方法

1.1 材料及处理

垫状卷柏采于云南省昆明市郊区，采集后进行培养箱培养（16 h 日照；25℃；相对湿度 20%）。采用垫状卷柏新鲜枝叶提取 DNA 和 cDNA；基因表达分析采用从岩石表层采集、植株根部土壤含量较低的垫状卷柏植株，对实验材料根部充分浇水后，进行自然干旱处理，在处理时间点进行新鲜幼嫩枝叶采集，分为 6 个处理组，分别为 0 h（根部充分着水，对照组）、自然干旱 2 h、自然干旱 4 h、自然干旱 12 h、自然干旱 24 h 及干旱 24 h 后复水 2 h 组（复水组在干旱 24 h 后立即给予充足水分）。各处理组选取 6 个长势一致的植株，3 个生物学重复，每个处理后的样本液氮速冻于-80 ℃保存。

1.2 方法

1.2.1 *SpLEA*1 基因 DNA 的克隆

垫状卷柏总 DNA 的提取按照小量植物（叶）总 DNA 抽提试剂盒（北京天根生化公司）的说明书提取。根据本实验室转录组测序结果，设计 *SpLEA*1 基因全长引物（表 1），以垫状卷柏 DNA 为模板，进行 PCR 扩增。PCR 反应体系：DNA 模板 2.0 μL、2×Taq PCR MasterMix 16 μL、上下游引物各 1.0 μL、ddH$_2$O 20 μL，终体积 40 μL。PCR 扩增程序：94 ℃，预变性，2 min；94 ℃，变性，30 s，57 ℃，退火，30 s，72 ℃，延伸，90 s，38 个循环；72 ℃，延伸，10 min；4 ℃保存。使用 1%琼脂糖凝胶电泳检测并使用凝胶回收试剂盒（OMEGA 公司）纯化 PCR 产物，连接到克隆载体 PMD18-T（宝日医生物技术）中，转化大肠杆菌 *E. coli* DH5α（天根生化生物公司），经菌液 PCR 鉴定后送生工生物工程股份有限公司进行测序。

1.2.2 *SpLEA*1 基因 cDNA 的克隆

垫状卷柏的总 RNA 按照总 RNA 提取试剂盒（OMEGA 公司）的说明书提取。参照逆转录试剂盒（全式金生物技术有限公司）的使用说明书方法，使用提取的 RNA 作为模板进行逆转录反应。以垫状卷柏 cDNA 第一链为模板，进行 PCR 扩增（引物、PCR 反应体系和程序同 1.2.1 相同），转化克隆、测序，获得 cDNA 全长。

1.2.3 *SpLEA*1 基因启动子的克隆

以垫状卷柏 DNA 为模板，结合 LEA1 基因序列设计启动子特异引物 *SpLEA*1Q1/2/3（表 1），分别作为第一轮、第二轮、第三轮特异引物，依次与随机引物组合，进行 HiTail-PCR 扩增，随机引物采用 Liu & Chen（2007）设计的 LAD1-1/2/3/4（表 1）。选择第三轮产物进行电泳检测，切取目的条带连接、转化进行克隆后送往测序。

1.2.4 生物信息学分析

利用 ExPASy-ProtParam（https://web.expasy.org/protparam/）分析氨基酸的理化性质；利用 Softberry（http://linux1.softberry.com/berry.）分析基因的结构信息；利用 NetPhos 3.1 Server（http://www.cbs.dtu.dk/services/NetPhos/）预测磷酸化位点；利用 Protscale（https://web.expasy.org/protscale/）分析蛋白亲/疏水性；利用 SOPMA（https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html）进行二级结构预测及分析；利用 Swiss-Model（https://swissmodel.expasy.org/interactive）通过同源建模建立三级模型；利用 DNAMAN9 软件获得多序列结构域比对图并分析遗传距离矩阵；利用 MEGA X 软件构建系统进化树；利用在线软件 PlantCARE(https://bioinformat-ics.psb.ugent.be/webtools/plantcare/heml/)对启动子顺式作用元件进行分析。

1.2.5 *SpLEA*1 的实时荧光定量 PCR 表达

利用实时荧光定量 PCR 技术分析垫状卷柏 *SpLEA*1 基因在干旱胁迫处理下的表达，根据获得的垫状卷柏 *SpLEA*1 基因序列设计一对 *SpLEA*1 定量 PCR 引物，以卷柏 Actin 为内参基因（表 1），以 cDNA 为模板，利用 TB Green® Premix Ex Taq™II说明书（Takara，RR820A）进行实时荧光定量 PCR 分析，每个样品 3 次重复以减小误差，反应程序：95 ℃，预变性，30 s；95 ℃，变性，5 s，60 ℃，退火，30 s，40 个循环。每个样品设有 3 次重复，采用 $2^{-\triangle\triangle Ct}$ 法。

表 1　本研究所用引物序列

Table 1 Primer sequences used in this study

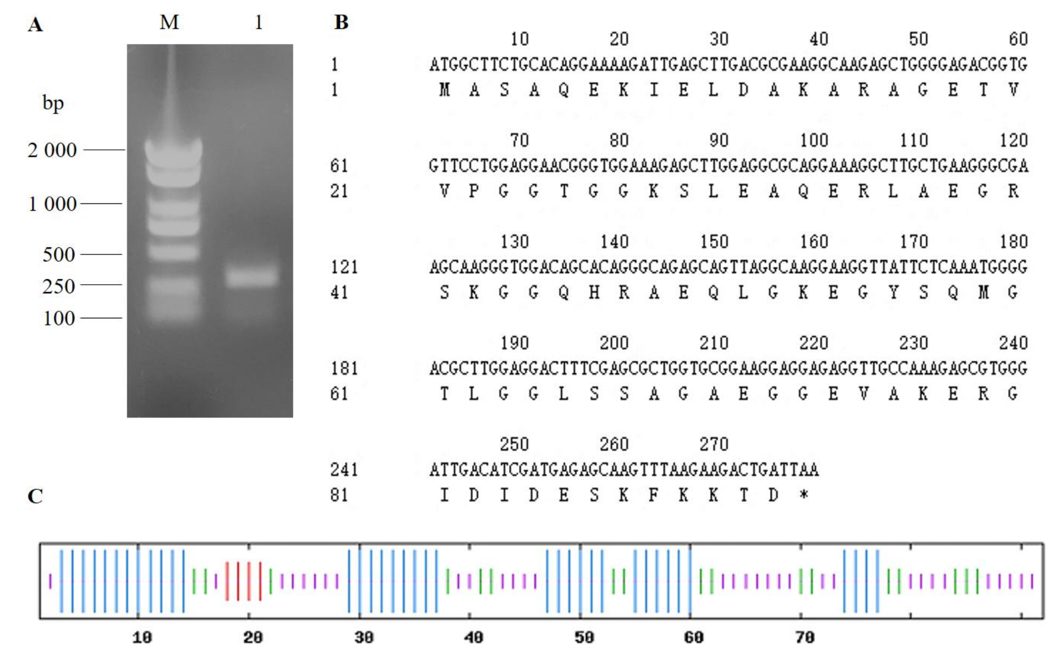
名称 Name	引物序列（5′→3′） Primer sequences (5′→3′)
*SpLEA*1-F	ATGGCTTCTGCACAGGAAAAG
*SpLEA*1-R	TTAATCAGTCTTCTTAAACTTGC
*SpLEA*1Q1	CTCATCGATGTCAATCCCACGC
*SpLEA*1Q2	CCTTCCTTGCCTAACTGCTCTG
*SpLEA*1Q3	CTTCAGCAAGCCTTTCCTGCG
QREJEM-1	CAGCACAGGGCAGAGCAGTTAG
QRTJEM-2	CCTCTCCTCCTTCCGCACCAG
SmAF	CCAACTGGGACGACATGGAGA
SmAR	CACCGCCTGAATAGCAACGT
LAD1-1	ACGATGGACTCCAGAGCGG CCGCVNVNNNGGAA
LAD1-2	ACGATGGACTCCAGAGCGG CCGCBNBNNNGGTT
LAD1-3	ACGATGGACTCCAGAGCGG CCGCVVNVNNNCCAA
LAD1-4	ACGATGGACTCCAGAGCGG CCGCBDNBNNNCGGT

2 结果与分析

2.1 *SpLEA*1 基因的克隆

以垫状卷柏总 DNA 为模板，结合引物 *SpLEA*1-F/R 进行 PCR 扩增，经测序鉴定获得全长 475 bp 的 *SpLEA*1 基因。以 cDNA 为模板，采用 *SpLEA*1-F/R 为引物，克隆获得 279 bp 的 *SpLEA*1 基因 cDNA 序列（图 1：A）。测序结果表明，*SpLEA*1 基因含有 1 个内含子（196 bp）和 2 个外显子（115 bp 和 164 bp），cDNA 长度为 279 bp，编码 92 个氨基酸（图 1：B）。根据 NCBI 的 PFAM 数据库查询，得到 SpLEA1 蛋白在 2~88 氨基酸位点含有保守结构域 LEA-5，PFAM 号是 PF00477，表明垫状卷柏 *SpLEA*1 基因属于 LEA1 家族。

通过 SOPMA（https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html）在线预测垫状卷柏 SpLEA1 蛋白质二级结构，由图 1：C 可知，α-螺旋占 40.22%，β-转角占 18.48%，无规则卷曲占 36.96%，延伸链占 4.35%，采用 Swiss-Model（https://swissmodel.expasy.org/interactive）在线工具进行同源建模，预测其三级结构。垫状卷柏 SpLEA1 蛋白主要由 α-螺旋和无规则卷曲构成，β-转角和延伸链占比较小。



A. *SpLEA*1 基因 cDNA 的扩增，其中 M 为 DL2000 DNA marker，1 为 *SpLEA*1 基因；**B.** *SpLEA*1 基因的 cDNA 序列及其编码的氨基酸序列；**C.** SpLEA1 蛋白二级结构预测。**蓝色.** α-螺旋；**绿色.** β-转角；**紫色.** 无规则卷曲；**红色.** 延伸链。

A. Amplification of *SpLEA*1 gene, where M is DL2000 DNA marker and 1 is *SpLEA*1 gene; **B.** The cDNA sequence of the *SpLEA*1 gene and the amino acid sequence encoded by it; **C.** Secondary structure prediction of SpLEA1 protein. **Blue**. α-helix; **Green**. β-turn; **Purple**. Random coil; **Red**. Extended chain.

图 1 *SpLEA*1 基因序列分析

Fig.1 Sequence analysis of *SpLEA*1 gene

2.2 *SpLEA*1 的生物信息学分析

通过 ExPASy-ProtParam（http://www.expasy.org/tools/protparam.html）预测分析，*SpLEA*1 基因编码蛋白质相对分子量为 9 491.46 Da，理论等电点 5.45，分子式为 $C_{397}H_{659}N_{123}O_{142}S_2$，带负电荷(Asp + Glu)的残基总数为 16，带正电荷(Arg + Lys)的残基总数为 14。不稳定指数为 28.72，该蛋白为稳定蛋白。SpLEA1 蛋白含 Thr（4.3%）、Lys（9.8%）、Gln（5.4%）、Gly（18.5%）、Glu（13.0%）亲水性氨基酸，还含有 Ala（12.0%）、Met（2.2%），Val（3.3%）、

Ile（3.3%）、Leu（6.5%）疏水性氨基酸。经计算，亲水性氨基酸占 51%，疏水性氨基酸占 27.4%。平均亲水指数为 -0.838。

通过 protscale（http://www.expasy.ch/tools/protscale.html）在线预测分析垫状卷柏 SpLEA1 蛋白的亲/疏水性，结果表明，小于 0 的亲水性氨基酸占多数，大于 0 的疏水性氨基酸只占少数，在第 68 位氨基酸有最大值为 0.892，该处疏水性最强，在第 43 位氨基酸有最小值为 -2.433，该处亲水性最强（图 2），说明 SpLEA1 蛋白属于亲水性蛋白。

利用 NetPhos 3.1 Server（http://www.cbs.dtu.dk/services/NetPhos/）在线预测发现 SpLEA1 蛋白发生磷酸化修饰的位点共有 10 个。其中丝氨酸有 6 个（Serine 位点为 3、29、57、66、67、86），而酪氨酸有 3 个（Tyrosine 位点为 19 和 25），苏氨酸仅有 1 个（Threonine 位点 56）。

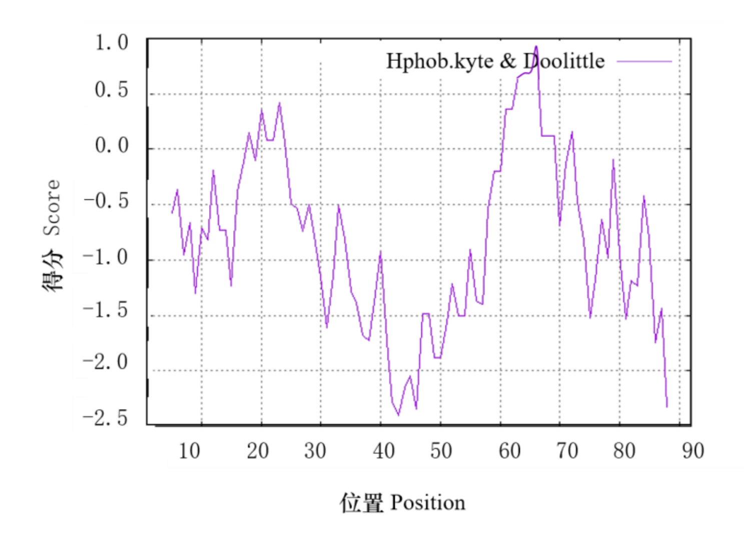


图 2　SpLEA1 蛋白质亲疏水性预测

Fig.2 Hydrophilicity prediction of SpLEA1 protein

利用 MEGAX 构建系统进化树，并使用 DNAMAN 进行遗传距离矩阵分析，发现垫状卷柏 SpLEA1 与鹰嘴豆(XP_004506729.1)和红三叶(PNX91110.1)蛋白同源性较高，聚为一支（图 3）。用 Blast 进行序列相似度分析，将垫状卷柏 SpLEA1 蛋白与鹰嘴豆、胡麻、黑麦等 15 种亲缘关系较近但物种不同的蛋白序列进行比对（图 4）。多重序列比对结果显示，SpLEA1 蛋白含有 LEA-5 保守结构域。在遗传距离矩阵中发现，红三叶和鹰嘴豆与垫状卷柏 SpLEA1 的遗传距离最近，为 0.272；荠菜与垫状卷柏 SpLEA1 的遗传距离最远，为 0.348。

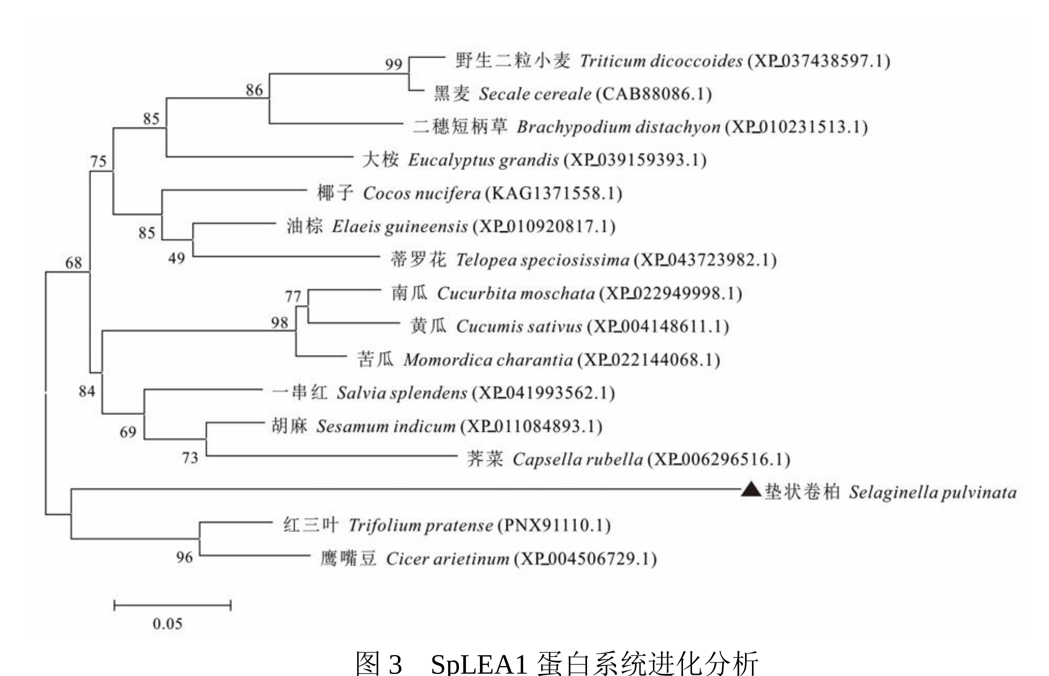


图 3　SpLEA1 蛋白系统进化分析

Fig.3 Phylogenetic analysis of the SpLEA1 protein

```
Sp  |                 .MASAQEKIEIIAKAFAGETVVPGGTGGKSLEAQERIAEG   39
Ss  | XP_041993562.1  MASQQQLRSEIIAKARQGETVVPGGTGGKSLEAQEHIAEG   40
Bd  | XP_010231513.1  .MASAQERSEIDRKAREGETVVPGGTGGHNLEAQEHIAEG   39
Cm  | XP_022949998.1  .MSSEQERCEIIAFARQGETVVPGGTGGKSLEAQEHIAEG   39
Eg  | XP_039159393.1  .MASRQEREEIDRFAKEGETVVPGGTGGKSLEAQEHIAEG   39
Cn  | KAG1371558.1    .MATQQERAEIIAKARQGETVVPGGTGGKSLEAQERIAEG   39
Egs | XP_010920817.1  .MATRQERAEIIAKARQGETVVPGGTGGHSLEAQEHIAEG   39
Tp  | PNX91110.1      .MASQQNRKEIDEKAKQGETVVPGGTGGKSLEAQEHIAEG   39
Si  | XP_011084893.1  .MASQQERKEIIAFAKQGETVVPGGTGGKSLEAQEHIAEG   39
Mc  | XP_022144068.1  .MSSEQERCEIIAFARQGETVVPGGTGGKSLEAQEHIAEG   39
Cr  | XP_006296516.1  .MASQQEKKQIDEFAKKGETVVPGGTGGKSLEAQQHIAEG   39
Ts  | XP_043723982.1  .MAAKQVRAEIIAKARQGETVIPGGTGGRSLEAQEHIAEG   39
Td  | XP_037438597.1  MASGQQERSQIDRKAREGETVVPGGTGGKSLEAQENIAEG   40
Ca  | XP_004506729.1  .MASQQNRQEIDEKARQGETVVPGGTGGKSLEAQEHIAEG   39
Cs  | XP_004148611.1  .MSSEQERGKIIAFARQGETVIPGGTGGKSLEAQEHIAEG   39
Sc  | CAB88086.1      MASGQQERSQIDRKAREGETVVPGGTGGKSLEAQENIAEG   40
Consensus              masqqer eldakarqgetvvpggtggksleaqehlaeg

Sp  |                 RSKGGQHRAEQIGKEGYSQMGTIGGISSAGAEGGEVAKER   79
Ss  | XP_041993562.1  RSRGGQTRREQIGTEGYQEMGRKGGISTGDKSGGERAGEE   80
Bd  | XP_010231513.1  RSRGGQTRKEQIGEEGYREMGRKGGISTMEESGGERAARE   79
Cm  | XP_022949998.1  RSRGGQTRKEQIGHEGYQEMGRKGGISNSGMPGGERAAEE   79
Eg  | XP_039159393.1  RSKGGQTRKEQIGTEGYQQMGRKGGISTGDESGGERTARE   79
Cn  | KAG1371558.1    RSRGGQMRREQIGTEGYQEMGRKGGISTMDESGGERAARE   79
Egs | XP_010920817.1  RSRGGQTRREQIGTEGYQEMGRKGGISTTDESGGERAARE   79
Tp  | PNX91110.1      RSKGGQTRKEQIGTEGYQQMGRKGGISTMEKSGGERAEEE   79
Si  | XP_011084893.1  RSRGGQTRREQIGTEGYQEMGRKGGISTGDKAGGERAEEE   79
Mc  | XP_022144068.1  RSRGGQTRKEQIGHEGYQQMGRKGGISNTGLSGGERAAEE   79
Cr  | XP_006296516.1  RSRGGQTRKEQIGTEGYQEMGRKGGISTGDKPGGENAEDE   79
Ts  | XP_043723982.1  RSRGGQTRKEQIGTEGYQEMGRKGGISTTEKSGGERAARE   79
Td  | XP_037438597.1  RSRGGQTRREQMGEEGYSQMGRKGGISTNDESGGERAARE   80
Ca  | XP_004506729.1  RSKGGQARKEQIGKEGYQEMGRKGGISTMDKSGEERAEEE   79
Cs  | XP_004148611.1  RSRGGQTRKEQIGHEGYQEMGRKGGISNTGMSGGERAAEE   79
Sc  | CAB88086.1      RSRGGQTRREQMGEEGYSEMGRKGGISTMDESGGERAARE   80
Consensus              rsrggqtrkeqlgtegyqemgrkgglst d sggeraaee

Sp  |                 GIDIDESKFKKTD.....                          92
Ss  | XP_041993562.1  GVEIDESKFRTRNT....                          94
Bd  | XP_010231513.1  GIEIDESKFKTKS.....                          92
Cm  | XP_022949998.1  GVEIDESKFRAK......                          91
Eg  | XP_039159393.1  GIEIDESKFRTKS.....                          92
Cn  | KAG1371558.1    GVDIDESKFRP.......                          90
Egs | XP_010920817.1  GIQIDESKFRT.......                          90
Tp  | PNX91110.1      GIDIDESKFKNQNK....                          93
Si  | XP_011084893.1  GIEIDESKFRTS......                          91
Mc  | XP_022144068.1  GVEIDESKFRTK......                          91
Cr  | XP_006296516.1  GVEIDESKFRTKT.....                          92
Ts  | XP_043723982.1  GIEIDESKFRKNT.....                          92
Td  | XP_037438597.1  GIDIDESKFKTKS.....                          93
Ca  | XP_004506729.1  GIDIDESKFKTGSGRNQN                          97
Cs  | XP_004148611.1  GVEIDESKFRNK......                          91
Sc  | CAB88086.1      GIDIDESKFKTKS.....                          93
Consensus              gieideskfrtk
```

Sp. 垫状卷柏；**Ss**. 一串红；**Bd**. 二穗短柄草；**Cm**. 南瓜；**Eg**. 大桉；**Cn**. 椰子；**Egs**. 油棕；**Tp**. 红三叶；**Si**. 胡麻；**Mc**. 苦瓜；**Cr**. 荠菜；**Ts**. 蒂罗花；**Td**. 野生二粒小麦；**Ca**. 鹰嘴豆；**Cs**. 黄瓜；**Sc**. 黑麦。
红线标注为保守结构域区间。

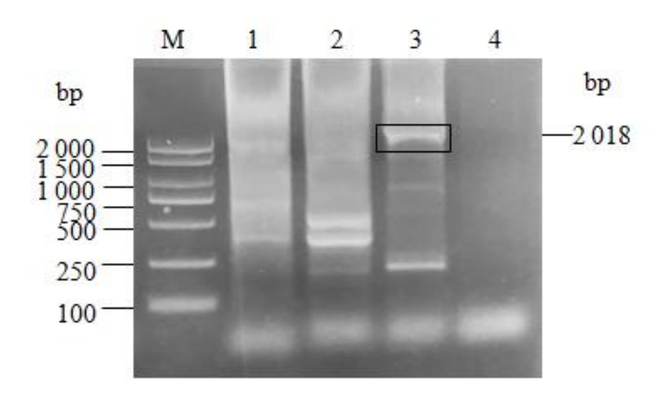
Sp. *Selaginella pulvinata*; **Ss**. *Salvia splendens*; **Bd**. *Brachypodium distachyon*; **Cm**. *Cucurbita moschata*; **Eg**. *Eucalyptus grandis*; **Cn**. *Cocos nucifera*; **Egs**. *Elaeis guineensis*; **Tp**. *Trifolium pratense*; **Si**. *Sesamum indicum*; **Mc**. *Momordica charantia*; **Cr**. *Capsella rubella*; **Ts**. *Telopea speciosissima*; **Td**. *Triticum dicoccoides*; **Ca**. *Cicer arietinum*; **Cs**. *Cucumis sativus*; **Sc**. *Secale cereale*. The red lines are marked as conservative structural domain intervals.

图 4　SpLEA1 蛋白与 15 种植物的多重序列比对

Fig.4 Comparison of SpLEA1 protein with homologous sequences of 15 plant species

2.3 *SpLEA*1 基因的启动子克隆与功能元件分析

以垫状卷柏 DNA 为模板，进行 HiTail-PCR 扩增 *SpLEA*1 基因启动子序列，选择第 3 泳道约为 2 000 bp 的条带回收产物（图 5），克隆获得 *SpLEA*1 起始密码子（ATG）上游 2 018 bp 的序列。使用 PlantCare 在线软件预测分析 *SpLEA*1 启动子中的顺式作用元件(图 6，表 2)，发现 *SpLEA*1 启动子区含有核心启动子 TATA-box 和启动子增强子常见顺式作用元件 CAAT-box；含有光反应元件（I-box 和 G-Box）和大量与非生物胁迫有关的功能性顺式作用元件，其中有激素类的脱落酸响应元件（ABRE）、茉莉酸甲酯响应元件（CGTCA-motif 和 TGACG-motif）、赤霉素响应元件（GARE-motif 和 P-box）、水杨酸响应元件（TCA-element）及辅酶响应元件（TGA-element），还包含与干旱胁迫有关的功能性元件 MYB 及参与干旱诱导的 MYB 结合点（MBS）。以上结果可推测 *SpLEA*1 基因的表达对垫状卷柏在干旱环境下的生存能力有显著影响。



M 为 DL2000 DNA marker，1、2、3、4 泳道为第三轮特异引物 *SpLEA*1Q3 和随机引物 LAD1-1/2/3/4 结合的 PCR 扩增产物。对第 3 泳道黑框内的产物进行回收。

M is DL2000 DNA marker, lanes 1, 2, 3, and 4 are specific primers *SpLEA*1Q3 and random primers combined with PCR amplification products. Recovery of the product in the black box of lane 3.

图 5　*SpLEA*1 基因启动子 HiTail-PCR 扩增

Fig.5 HiTail-PCR amplification of *SpLEA*1 gene promoter

-2 018 GAGGGAACGCCGCGCGCCCAGGCCAACACATTCTCCTCTTGATCACCTACAAGCATTCCCTCAGGAGATCCGGGAGGAGTTTCTATGCCGAGGCGGTTGTTCTTCCGG

TGA-element MYB/MBS

-1 910 CTGTTCTTTCAGTGCAGATGGAAT*CCAAT*GGTCAGAACTCTTTACGGGG**ATATAT**TTCACA*CAAAT*CTTTCACCTTCCAGTCGTTCTTGGATTTAGATTTCACCGTACAAC

MYB/MBS

-1 799 AGCTCAAGAAACTCTTGCTAGAGTT*CAAT*GCGGATTGTTCTGATCATTACCAGAGAGTA*CAAAT*TGGTCACTGTTTCATGCACTCGAG*CAAT*CCAAAGACAGTGAGCTC

-1 690 CAAAATCTTGTCCGTGAGCAGCTCTAAACCTCTTATTAG**TATA**TGTTTCGCCACCTCGGAAGAATCA*CAAT*GCCAACTTGTTCAGACAGAGAGTGACT**TATA**TGAGACCA

MYC ABRE ARE

-1 579 CTTTTGATAAGTATGTCCTGGTGAAAGT*CAAT*GCACCACAAAATGCATGTGATGGTATGAAATTTGTATCCATTGAAGGACAACGTGGTGTTCGTATGGTTTCTTCCTTCA

TGACG-motif

-1 470 TGGCATATGTTTCTGACGT**TATA**AAACGCTCAACTCCACCTGTACAAGAAATTTGTGACATTCTTCAGGAAGCTCGAAGCCATTCGTTGCCCGAGACGTCGATCGCGGA

MYC

-1 359 ACCTCTTCTTACGTTTCGAGGTTT**TATA**GTTCTTGTTAAACCGA*CAAATG*ATTCATCGTTGGAGCTAGAAGTTAAGGATTATTACGGAACA*CAAT*CGGTTTCTGTCCTTGT

-1 249 AGATGGATCCAAGGTCATTGTCCCTGTAAACTTGCGACCTGGAACCATCGCATCATTTCATCGAACTTTATTAAGAAGATGTTCAGATGGAAATATTGTGCTCCAGACAA

-1 139 TTCCTACGAGTGTAATTACTGTTCACGCCGTTTACAAGCTGAAGAAGACA**TATA**GTAAATTATTGGATAAAATAAATGAAACTCAGTACATGGATATGAATCTAGAACCG

ABRE

-1 028 AAGAAATGGTGC**TATATAT**CATGCTTGTCAAACACGTCGTC**TATA**ACGGGAATAGGCCTTCATTGCAGAGTCTTTGC**TATA**TTAGAACTTGTGTTTAAGTGGCTGACAAGT

P-box MYB/MBS

-917 CCAAAAGGAAGTACTGCACATAATCCTCTACAAAATATCGACTTTTCTGCTGGTTGCTTTCTTATTGATGATGGATCTGGTGTGGCCGAGTGCTGTGTGTATGGTGCATCA

ABRE/ TGA-element ARE

-808 GCCGTAAGTGGGTTTCTACAGCTAAGTTCGCATTCCGTGCCAGGAGGCATG**TATA**AGGACACAGTCCAGGAGTTTTGTACGGTCGTTCAAAACCATGGTCATGTAGTTT

MYC

-700 ACAGAAAAGAATACTCTGATGATTGGACACAACAA*CAAT*GTCAAGTTGTTCGAAATGATAAGAGTCTGAGTGAGACAGAGCAAGCCATAGTGGAGTCATTTGTGGCTA

G-Box /ABRE

-589 GAACATCTCGCACTGTACCCCTGGTATCATTTTGTATGCGTGGGTGTGTACTCGTTGACTTAGC**TATA**CAAGTTTAACGTGTAATGTGTATCAGGATAGTTTTGAGAAATA

MYB/MBS ARE

-481 TCAGTTGACACACACACACACATTTTGCAGGTTTTGGTTTGTGATTTACTGAGGGGCAGTGAAAACAGTAAAAGAGTGAGGTTAGTGGCAACTTCTGTGAAGACACTT

MYB/MBS/ GARE-motif LTR MYC

-371 TGCTACACAACAGAAGCCGAAAGATTGAGCAGGGAGCTGGAGCTTACGTTAAAGATCAAACATGTATGATCTATTTAACAGCTTCAGGTAGTGTATTCTATCATTTGATA

-261 TTCATACCAGCATACATTGTGTGTTGAGAGAGAGCATACTTGTAGACACCATTCCTAATGGCTTACGATGTTAG**TATA**GTGAGTAAAGTACGTAAGTAGTTACACACACA

ABRE

-150 GGACTTAATTGTTGGGGTAATTGTGTTTGTCAGGGTATTAACATTCTTTAAACGCGGTGTCAGCTTAG*CAAT*ATGACACG**TATA**CTGCTTGCATGCAAACATCAGTGCTAT

GARE-motif

-110 A**TATA**CATAATATCTGTTGGTGAGTTTTCAGATATC**TATA**

方框所框为顺式作用元件；黑色加粗的序列为 TATA-box；斜体的序列为 CAAT-box。
The box is *cis*-element; the black and bold nucleotide sequence is TATA-box; The sequence in italics is CAAT-box.

图 6 *SpLEA*1 基因启动子的序列及部分顺式作用元件

Fig.6 Sequence and partial *cis*-acting elements of the *SpLEA*1 gene promoter

表 2 *SpLEA*1 启动子部分顺式作用元件的预测

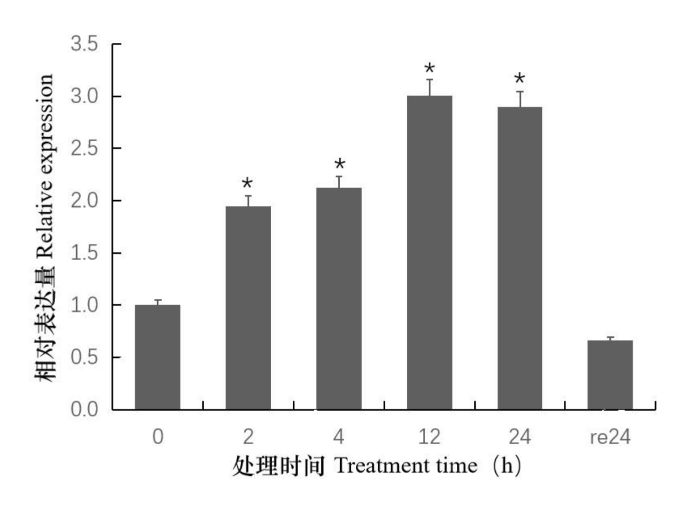
Table 2 Prediction of some *cis*-acting elements of *SpLEA*1 promoter

顺式作用元件 *Cis*-acting element	核心序列 Core sequence	功能 Function
TATA-box	TATA/ATATAT	转录起始点-30 附近的核心启动子元件 Core promoter element around -30 of transcription start
CAAT-box	CAAT/CCAAT/CAAAT	启动子和增强子区域的共同顺式作用元素 Common *cis*-acting element in promoter and enhancer regions
LTR	CCGAAA	参与低温反应性的顺式作用元素 *Cis*-acting element involved in low-temperature responsiveness
MYB	CAACAG/CAACCA	
MBS	CAACTG	参与干旱诱导的 MYB 结合点 MYB binding site involved in drought-inductility
I-box	GTATAAGGCC	光反应元件的一部分 A part of light-responsive element
GARE-motif	TCTGTTG	赤霉素反应元件 Gibberellin-responsive element
P-box	CCTTTTG	赤霉素反应元件 Gibberellin-responsive element
G-Box	CACGTT	参与光反应性的顺式调节元件 *Cis*-acting regulatory element involved in light responsiveness
ARE	AAACCA	厌氧诱导所必需的顺式作用调节元件 *Cis*-acting regulatory element essential for the anaerobic induction
TGACG-motif	TGACG	参与 MeJA 反应性的顺式作用调节元件 *Cis*-acting regulatory element involved in the MeJA responsiveness
TCA-element	TCAGAAGAGG	参与水杨酸反应性的顺式作用元素

		Cis-acting element involved in salicylic acid responsiveness
TGA-element	AACGAC	辅酶反应元件 Auxin-responsive element
ACE	GACACGTATG	参与光反应性的顺式作用元素
		Cis-acting element involved in light responsiveness
ABRE	ACGTG/TACGGTC/GC	参与脱落酸反应性的顺式作用元素
	AACGTGTC	*Cis*-acting element involved in the abscisic acid responsiveness
MYC	CATGTG/CATTTG	
CGTCA-motif	CGTCA	参与 MeJA 反应性的顺式作用调节元件
		Cis-acting regulatory element involved in the MeJA responsiveness

2.4 *SpLEA*1 基因实时荧光定量分析

利用 qRT-PCR 分析检测垫状卷柏 *SpLEA*1 在干旱胁迫下的表达模式。由图 7 可知，在自然干旱时间延长的过程中，*SpLEA*1 呈现上调表达趋势，在干旱 12 h 时达到峰值，后呈现下降趋势，在干旱 24 h 后给予复水处理 2 h，*SpLEA*1 的表达量显著下降。以上结果可以推测，*SpLEA*1 参与了垫状卷柏干旱胁迫的响应。



*表示与 0 h 处理相比差异显著（P<0.05）。0、2、4、12、24、re24 为植物材料自然干旱处理时间（h），其中 re24 代表干旱 24 h 后复水 2 h 的样品数据。

* means significant differences compared to 0 h treatment (P<0.05). 0, 2, 4, 12, 24, re24 are the dehydration treatment times of plant materials (h), where re24 h represents the data of samples rehydrated for 2 h after 24 h of dehydration.

图 7 *SpLEA*1 基因干旱胁迫下的表达

Fig.7 Expression of *SpLEA*1 gene under drought stress

3 讨论与结论

在本研究中 SpLEA1 蛋白保守结构域为 LEA-5（PF00477），属于 LEA1 家族。与 NCBI 中已收录的 15 个 LEA 蛋白进行序列对比，发现有 2 个保守序列在 N 端和 C 端出现，这与 Battaglia 等（2008）的报道一致，在同源进化关系的分析中，没有检索到与 SpLEA1 蛋白相似度高的同属植物，但发现垫状卷柏 SpLEA1 与鹰嘴豆和红三叶的 Lea-5 蛋白同源性较高，而在研究鹰嘴豆中 *CarLEA*793 和 *CarLEA*4 基因在干旱胁迫下保护植物细胞的分子机制，推进了 LEA 蛋白的耐旱性研究（顾汉燕，2010）。 对大豆中 LEA1 蛋白结构研究表明 LEA1 蛋白保守基序 Em-C 和 Em-2M 多肽在不同环境中的结构及聚集行为会发生改变，但都主要以无规则结构形式存在，这可能与 Em 蛋白在不同环境中的结构特点，及其重要区域在整个蛋白中所起的作用有关（薛蓉等，2012）。在邹永东（2011）的研究中可以发现 LEA1 蛋白

在应对不同环境时可以形成不同的空间结构，这些特殊的空间结构与 α-螺旋相互作用，对细胞中的酶活性与蛋白质结构起到了保护作用。利用在线软件 SOPMA 和 Swiss-Model 对 SpLEA1 蛋白的二级结构和三级结构进行分析预测，结果表明 SpLEA1 蛋白主要结构是 α-螺旋和无规则卷曲，与 LEA1 家族基因的结构特点一致，推测 SpLEA1 蛋白通过形成 α-螺旋与无规则结构参与对垫状卷柏的干旱响应。而在实时荧光定量结果分析中，可以看出 *SpLEA*1 基因在垫状卷柏干旱处理后出现表达上调的趋势，在干旱 12 h 时达到峰值，而在复水处理后表达量明显下降，反映了 SpLEA1 蛋白在受到干旱胁迫时会被诱导表达，这可能是垫状卷柏高耐旱性的原因之一。垫状卷柏 *SpLEA*1 基因的时空表达差异为后续过表达功能研究提供了基础。

本研究采用 HiTail-PCR 技术扩增 *SpLEA*1 基因启动子序列，利用在线软件 PlantCare 分析垫状卷柏 *SpLEA*1 启动子中的顺式作用元件，分析 *SpLEA* 1 基因的转录表达水平。发现了核心元件 TATA-box 和 CAAT-box，表明 *SpLEA*1 基因功能可以稳定表达，对 *SpLEA*1 启动子功能元件进行分析，结果表明启动子区中含有大量的诱导型启动子，包含 5 类激素诱导表达元件和非生物胁迫诱导表达元件，其中与水分胁迫有关的元件有 7 个 ABRE、4 个 MYB 和 4 个 MYC，与实时荧光定量实验中垫状卷柏在干旱处理后 *SpLEA*1 基因表达上调的结果相一致，反映了 *SpLEA*1 基因参与了干旱胁迫的应答响应。关于 *SpLEA*1 基因抗旱性的作用机制有待进一步研究。

本研究从垫状卷柏中克隆获得 *SpLEA*1 基因 cDNA 全长序列，通过对其蛋白结构分析得出 SpLEA1 蛋白为亲水性稳定蛋白。通过对 *SpLEA*1 基因启动子的分离克隆和顺式作用元件分析，可以推测 *SpLEA*1 基因在垫状卷柏耐旱机制中起作用，通过实时荧光定量表达实验可以观测到在干旱环境中 *SpLEA*1 基因的高表达特性。因此，推测 *SpLEA*1 基因与垫状卷柏的高耐旱性有关，参与了垫状卷柏在干旱胁迫下的表达调控。在后期的实验中，可以构建 *SpLEA*1 基因在酵母或拟南芥中的高表达载体，验证 *SpLEA*1 基因在原核和真核生物中的表达情况。对本研究的进一步探索能为植物抗旱基础研究领域提供一定依据，提高园林园艺经济观赏植物在干旱环境下的生存率。

参考文献：

BATTAGLIA M, OLVERA-CARRILLO Y, GARCIARRUBIO A, et al., 2008. The enigmatic LEA proteins and other hydrophilins[J]. Plant Physiol, 148(1): 6-24.

DURE L, GREENWAY SC, GALAU GA, 1981. Developmental biochemistry of cotton seed embryogenesis and germination: changing mRNA populations as shown *in vitro* and *in vivo* protein synthesis[J]. Biochemistry, 20: 4162-4168 .

D'URZO N, MARTINELLI M, NENCI C, et al., 2013. High-level intracellular expression of heterologous proteins in *Brevibacillus choshinensis SP*3 under the control of a xylose inducible promoter[J]. Microb Cell Fact, 12(1): 12.

GU HY, 2010. Cloning, expression, and subcellular localization of two LEA protein genes, *CarLEA793* and *CarLEA4* in chickpea(*Cicer arietinum* L.) [D]. Nanjing: Nanjing Agricultural University: 83. [顾汉燕，2010. 鹰嘴豆胚胎晚期富集蛋白基因 CarLEA793 和 CarLEA4 的克隆、表达分析及亚细胞定位[D]. 南京：南京农业大学：83.]

HUNAULT G, JASPARD E, 2010. LEAPdb: a database for the late embryogenesis abundant proteins[J]. BMC Genomics, 11 (1): 221.

JASPARD E, MACHEREL D, HUNAULT G, et al., 2012. Computational and statistical analyses of amino acid usage and physico-chemical properties of the twelve late embryogenesis abundant protein classes[J]. PLoS ONE, 7 (5): e36968.

LI X, 2016. Molecular evolutionary research of the plant LEA gene family and preliminary functional analysis of an atypical LEA gene in *Arabidopsis*[D]. Zhenjiang: Jiangsu University: 84. [李翔，2016. 植物 LEA 基因家族的分子进化研究和拟南芥非典型 LEA 基因功能的初步分析[D]. 镇江：江苏大学：84.]

LIU YG, CHEN Y, 2007. High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences[J]. Biotechniques, 43(5):649.

LIU ZB, HAGEN G, GUILFOYLE TJ, 1997. A G-Box-Binding protein from soybean binds to the E1 auxin-response element in the soybean GH3 promoter and contains a proline-rich repression domain[J]. Plant Physiol, 115(2): 397-407.

MEI YQ, 2018. Effect of *DAD* gene promoter sequence variation on upland cotton domestication[D]. Qufu: Qufu Normal University: 73. [梅玉芹，2018. 陆地棉 DAD 基因启动子序列变异对其驯化的影响[D]. 曲阜：曲阜师范大学：73.]

SAUCEDO AL, HERNÁNDEZ-DOMÍNGUEZ EE, DE LUNA-VALDEZ LA, et al., 2017. Insights on structure and function of a late embryogenesis abundant protein from *Amaranthus cruentus*: An intrinsically disordered protein involved in protection against desiccation, oxidant conditions, and osmotic stress[J]. Front Plant Sci, 8: 497.

SILVEIRA V, SANTA-CATARINA C, BALBUENA TS, et al., 2008. Endogenous abscisic acid and protein contents during seed development of *Araucaria angustifolia*[J]. Biol Plantarum, 52: 101-104.

STACY RA, ESPELUND M, SBE-LARSSEN S, et al., 1995. Evolution of the Group 1 late embryogenesis abundant (Lea) genes: analysis of the Lea *B*19 gene family in barley[J]. Plant Mol Biol, 28(6): 1039-54.

WANG Z, ZHAO L, LI WB, 2011. Advances in the study of inducible promoters in plants[J]. Soybean Sci Technol, (3): 5-9. [王志，赵琳，李文滨，2011. 植物诱导型启动子的研究进展[J]. 大豆科技，(3)：5-9.]

WISE MJ, 2003. LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles[J]. BMC Bioinformatics, 4:52.

WU ZY, 2004. FRPS[M]. Beijing: Science Press:100. [吴征镒，2004. 中国植物志[M]. 北京：中国科学院科学出版社: 100.]

XIANG DJ, MAN LL, ZHANG CL, et al., 2018. A new Em-like protein from *Lactuca sativa*, *LsEm*1, enhances drought and salt stress tolerance in Escherichia coli and rice[J]. Protoplasma, 255(4):1089-1106.

XUE R, ZOU YD, ZHENG YZ, et al., 2012. Structures and assembly of conserved domains from soybean late embryogenesis abundant protein Em[J]. Anal Chem, 40(4): 563-568. [薛蓉，邹永东，郑易之，等，2012. 大豆 Em(LEA1)蛋白保守结构域的结构和聚集特性[J]. 分析化学，40(4)：563-568.]

YANG RJ, BAI JR, LI R, et al., 2018. Research progress of the inducible promoters in plant genetic engineering[J]. Shanxi Agric Sci, 46(2): 292-298. [杨瑞娟，白建荣，李锐，等，2018. 诱导型启动子在植物基因工程中的研究进展[J]. 山西农业科学，46(2)：292-298.]

ZHANG XY, HE H, LI Z, et al., 2019. Functional analysis of the *A7-FT* promoter in *Brassica napus* L[J]. ACTA Agric Boreal-Sin, 34(02): 59-65. [张曦予，贺慧，李祯，等，2019. 甘蓝型油菜 A7-FT 启动子的功能分析[J]. 华北农学报，34(02)：59-65.]

ZHENG J, SU H, LIN R, et al., 2019. Isolation and characterization of an atypical LEA gene (IpLEA) from *Ipomoea pes-caprae* conferring salt/drought and oxidative stress tolerance[J]. Sci

Rep, 9(1): 14838.

ZOU YD, 2011. Molecular mechanism for the enhanced salt tolerance of *Arabidopsis thaliana*, *Saccharomyces cerevisiae* and *Eacherichia coli* by soybean EM(LEA1) protein[D]. Changchun: Northeast Normal University: 99. [邹永东，2011. 大豆 Em（LEA1）蛋白提高拟南芥、酵母和大肠杆菌耐盐性的分子机理[D]. 长春：东北师范大学：99.]